From: Quernemoen, John M
Sent:
To: Herrera, Ricardo -NOTES; Johnson, David Wayne -NOTES; Meszaros, John H.; 'Mr. Toshiro Sugaya'; Brunner, Rick -NOTES; Cook, Russ; Strum, Steven D. -NOTES; Lyes, John A. -NOTES
Cc: Buggert, Jerrold E; Van Oosterom, Hendrikus; Freeman, Rick D
Subject: NT Sizer Update

Note: You may have already received the following message. If so, ignore it.
****
Attached find an updated version of the NT Sizer. The two files, NTSizer_280.exe and NT Sizer User Guide.exe, are WinZip compressed files.

NTSizer_280.exe contains the following files:
- Excel add-in file NTSizer_280.xla which should be loaded to use the sizing tool
- *.xlt files which are workbook templates used dynamically by the sizer
- Sample System Sizing.xls which is a sample output from a sizing

NT Sizer User Guide.exe contains the user guide in pdf (Adobe Acrobat) format

This version requires Office 97 to run. Also, it has been modified to run on the Japanese version of Excel as well.

Updates since the class are:
- Support for SQL Server 7 and Oracle 8 OLTP applications
- Comparators for both TPC-C and TPC-D, updated
- Inclusion of NIC Sizing

Updates since the class are:
- Providing a more detailed manner of determining mass storage capacity requirements
- Detailed mass storage capacity calculations for SQL Server 6.5, 7.0, snf Oracle 8 databases

If you have any problems installing and running the sizer, please let me know.

Also, note that our Web page has several additions and enhancements.

Regards

Mike Quernemoen


NT Sizer User Guide.exe


NTSizer_280.exe

EXHIBIT 1

Appendix A:
Unisys Enterprise NT Sizer Description and User Guide
Version 2.80

EXHIBIT 2

Unisys
NT Performance Services

Unisys Enterprise NT Sizer Description & User Guide Version 2.80

System Analysis, Modeling, and Measurement
Unisys Corporation
Roseville, MN

Table of Contents

1. Introduction

The Unisys Enterprise NT Sizer provides system configuration recommendations for Windows NT Database Server systems, based on customer workload requirements for applications running relational database-management systems (DBMS) in a client/server computing environment. The Sizer is a Windows application that runs within a Microsoft Excel application framework. Within this framework, the Sizer was designed to offer a user friendly interface to accomplish complex tasks.

This document describes the NT Sizer in the following manner:

- Its current capabilities are summarized;
- A user guide is provided;
- The methodology behind the Sizer is described.

2. NT Sizer Features

The sizer is a Windows application that runs as a Microsoft Excel add-in. The sizer requires the Excel 97 version to run. The sizer was developed on a system using a 17" monitor with 1024 x 768 resolution and small fonts; although the sizer has been tested with smaller monitors and coarser resolution and font settings, it is possible that some dialogs may not be as easy to use in these cases.

This tool is intended for use by analysts who will assist customers in selecting an NT system configuration that will meet the customers' DBMS application and workload needs. Thus, it is designed to interact with the user to define the database, the application, the application workload and growth. The current version of the Sizer has the following features.

2.1 Comparator

The TPC Comparator calculates ratios of performance and price/performance for published multi-vendor Transaction Processing Council Benchmark C (TPC-C) and Transaction Processing Council Benchmark D (TPC-D) measurements. The cases cover the UNIX and Windows NT 4.0 operating systems, as well as several different relational database management systems.

The data used by the Comparator is first downloaded from the TPC WEB site http://www.tpc.org and then processed to provide a user friendly interface to make quick comparisons of performance and price/performance.

2.2 Size Pre-defined Workloads

The Unisys Enterprise NT Sizer estimates system configuration requirements for Online Transaction Processing (OLTP) applications and workloads. These workloads may be pre-defined or user defined. Included in the set of pre-defined workloads is the TPC-C benchmark. Other specific, pre-defined applications will be included in future releases of the Sizer.

2.3 Size User Defined Workloads

The Sizer also allows the user to define applications and workloads according to the business case of the potential customer. The data gathered via the various dialogs for an OLTP system consists of:

- Size of the database and its various indexes
- A definition of the composition of each transaction in the application, e.g., number of SQL inserts, deletes, updates, and selects
- The expected transaction rate, measured in transactions per second (TPS), by transaction
- The processor, network interface card, and mass storage device utilization requirements

From this data the following system sizing information can be estimated:

- Processor requirement, and expected processor utilization
- Memory requirement
- Mass Storage requirement
- Number of users supported
- Network Interface Card requirement and expected utilization

Currently, estimates are available for applications using the SQL Server 6.5, SQL Server 7.0, and Oracle 8.04 DBMS. The systems supported are the XR/6 and QS/2 Enterprise systems.

Estimates given by the sizer are from the point of view of the NT Server only. Client configuration estimates are currently beyond the scope of this tool.

2.4 *Mass Storage Sizing Only*

This feature, which is already incorporated into the process of system sizing for a user defined workload, allows the user to address mass storage requirements only. Two types of sizing are available. If the user has a very vague notion of the database definition, a rough estimation capability is provided, and is based on the input of only six parameters. If the user has more specific information about each table's size and the characteristics of its indexes, then a more detailed mass storage requirement estimation capability is available. In either case dialogs are provided to change some parameter settings which will be more suitable to the user's requirements.

3. User Guide

The NT Sizer installation diskette contains a single self-extracting zip file. To unzip, simply double-click on the file (**NTSizer_280.exe**) in Windows Explorer. A WinZip dialog box will appear, asking you to enter the path name where you would like the NT Sizer files placed. After entering the path name, click the **Unzip** button. When the unzip process is complete, a message box indicating that three files were unzipped will be displayed. Click **OK**, then click **Close** on the WinZip dialog box. The NT Sizer installation is now complete.

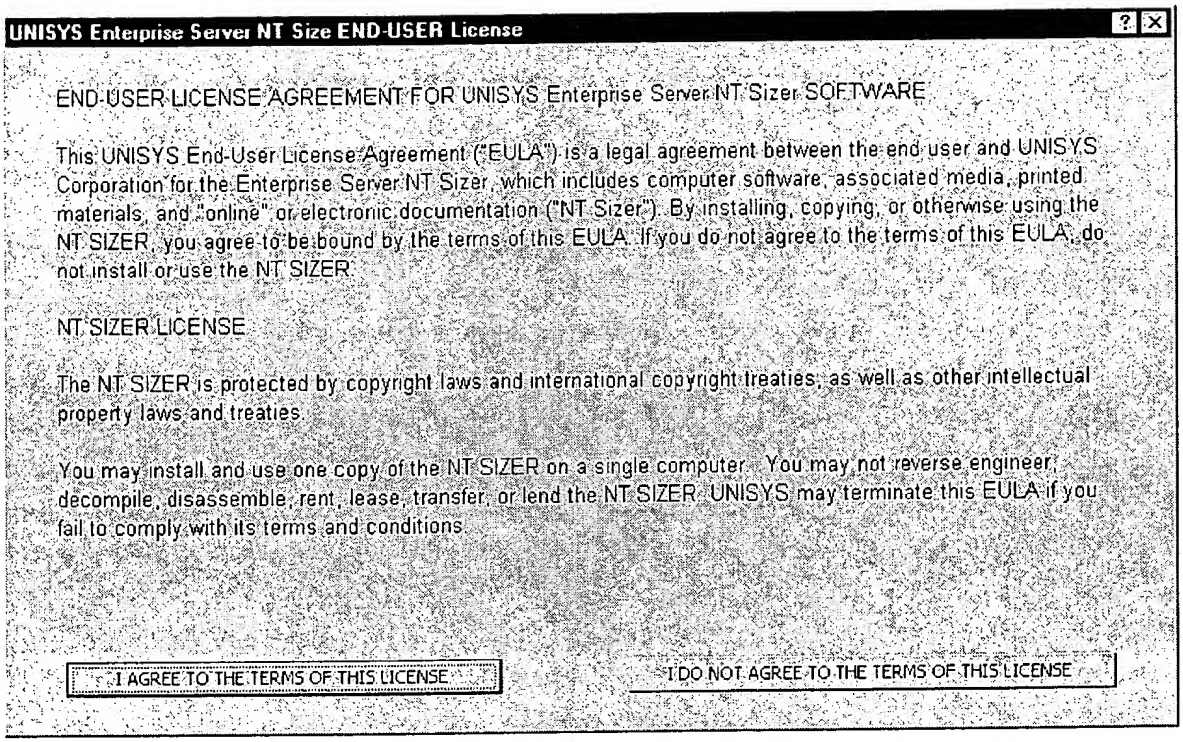
Five files are required to run the NT Sizer. One is basically an executable which runs under the Excel framework. The other two support the functionality of the Sizer. More specifically, these files are the following:

- **NTSizer_280.xla**, an Excel add-in file which contains the menu items, procedures and database required to perform a sizing;
- Three templates which contain preformatted worksheets used by the Sizer to store information frot he user:
  - **UserBk.xlt**
  - **Mass Storage Estimate.xlt**
  - **Mass Storage Estimate Detailed.xlt**
- **Lic.xls,** a password-protected workbook containing access information for the Sizer.

All five files must reside in the same directory.

To start the Sizer, double-click on the add-in file, **NTSizer_280.xla** from Windows Explorer, or start Excel and open the file using the **File Open** dialog box. The first time you use the Sizer, you will see the following license agreement.

This license agreement will appear only once. Accepting the license agreement completes the installation. Declining causes the Sizer to unload.



Each time the Sizer starts, a message box appears and provides the following information:

- Brief summary of the Sizer's functionality; the Sizer contains primarily two tools: the **Enterprise Server NT Sizer for Windows NT Applications**, and the **TPC Comparator**. The Enterprise NT Sizer estimates system configuration requirements based on the database size and transaction workload. The Comparator calculates ratios of published TPC-C and TPC-D results across several vendors, operating systems, and relational database systems.
- Navigate through the sizer's functionality via the menu bars specially created for use with the sizer.
- The Sizer's expiration date. The Sizer's **Help** menu bar directs you to contacts so that the Sizer's status can be upgraded.

Click **OK** to close the message box.

When the Sizer is opened, a copy of the workbook template UserBk.xlt is loaded as UserBk*n*.xls where the value of *n* is dynamically assigned by the Excel program. This workbook acts as a storage area for system sizing results and input data where the writing of the information to this workbook is controlled by the Sizer. One may view how the data is stored by executing the macro that comes with Userbk.xls, **MakeAllSheetsVisisble**. This is done by selecting **Tools, Macros.** Note that Userbk*n*.xls contains several worksheets. The meaning and use of these worksheets will be discussed later.

3.1 The Sizer Menu

The Sizer is operated by making selections from each of the three menu options added to the Excel menu bar: **Sizer Menu, Workload,** and **Sizer Help**. Note that all other Excel functionality is still available. Primary sizer functionality is initiated via the **Sizer Menu.**
Selecting the **Sizer Menu** gives the following options:

- **Comparator**
  - **TPCC Comparator**
  - **TPCD Comparator**
- **OLTP System Sizing**
  - **TPCC Workload**
  - **User Defined Workload**
- **Mass Storage Sizing**
- **Close Sizer**
- **Exit Excel**

Selecting **Close Sizer** will cause the sizer add-in to unload from memory and the sizer menu to be removed. Selecting **Exit Excel** is the same as selecting **File, Exit**.

The remaining subsections describe the capabilities associated with the **Comparator, OLTP System Sizing,** and **Mass Storage Sizing** features.

3.2 Using the Comparator

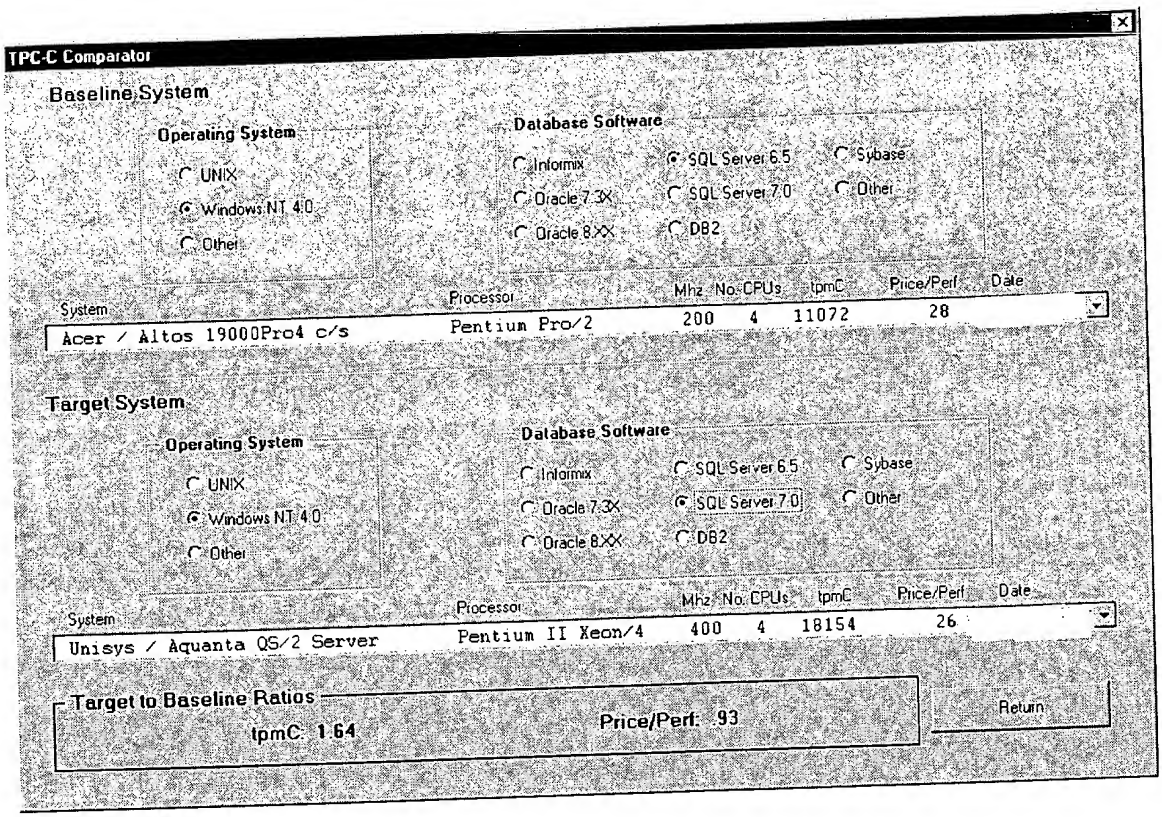
The TPC Comparator calculates ratios of performance and price/performance for published multi-vendor TPC-C and TPC-D measurements. The cases cover the UNIX and Windows NT 4.0 operating systems, as well as several different relational database management systems.

The data used by the Comparator is first downloaded from the TPC WEB site http://www.tpc.org and then processed to provide a user friendly interface to make quick comparisons of performance and price/performance. We also note that all of the ratings published at the TPC site are based on 100% processor utilization.

Two tools are used to make comparisons. One is a comparison of metrics for the TPC-C benchmark, and the other, TPC-D benchmark.

3.2.1 TPC-C Comparator

To start the TPC-C Comparator, select **Comparator , TPCC Comparator** from the **Sizer Menu**. A dialog box, like the one below, will be displayed where you can choose a number of values for a baseline and target system. The Comparator will then calculate the ratio of tpmC's (transactions per minute) and $ / tpmC (price per tpmC) of the target system to the baseline system.



The two standard metrics used for comparison for the TPC-C benchmark are tpmC and $/tpmC. tpmC stands for transactions-per-minute-C; it is a measure of "business throughput" and represents the number of orders processed per minute. Five different transaction types are used to simulate the business activity of processing an order: New-Order, Payment, Order-Status, Delivery, Stock-Level. tpmC is the number of New Order transactions processed per minute. New Order transactions account for approximately 45% of the total TPC-C transaction workload. More information on TPC-C can be found on both the Performance Lab and TPC WEB sites [8,9].

To run the TPC-C Comparator dialog, first select the characteristics of the baseline system, i.e., the system to which you want to compare performance. The first choice is that of the **Operating System**, with the choices being **Unix** or **Windows NT 4.0**, or **Other**. The **Other** choice includes, for example, the IBM AS/400 Operating System. The second choice is that of the **Database Software**, where the following relational database management systems (RDBMS) are available: **DB2**, **Informix**, **Oracle 7.3** and **8.0**, **SQL Server 6.5** and **7.09**, **Sybase**, or **Other**. The **Other** choice includes proprietary RDBMS'. Once the operating system and RDBMS have been selected, a drop down list then shows which systems were tested and results subsequently published with the TPC-C benchmark measurement. Each line on the drop down list gives a summary system description plus the value of the New-Order rate (tpmC) and the corresponding price/performance estimate for the system measured. If the operating system and RDBMS choices are such that no values are available (e.g., SQL Server 6.5 on Unix), the drop-down list will show **None**.

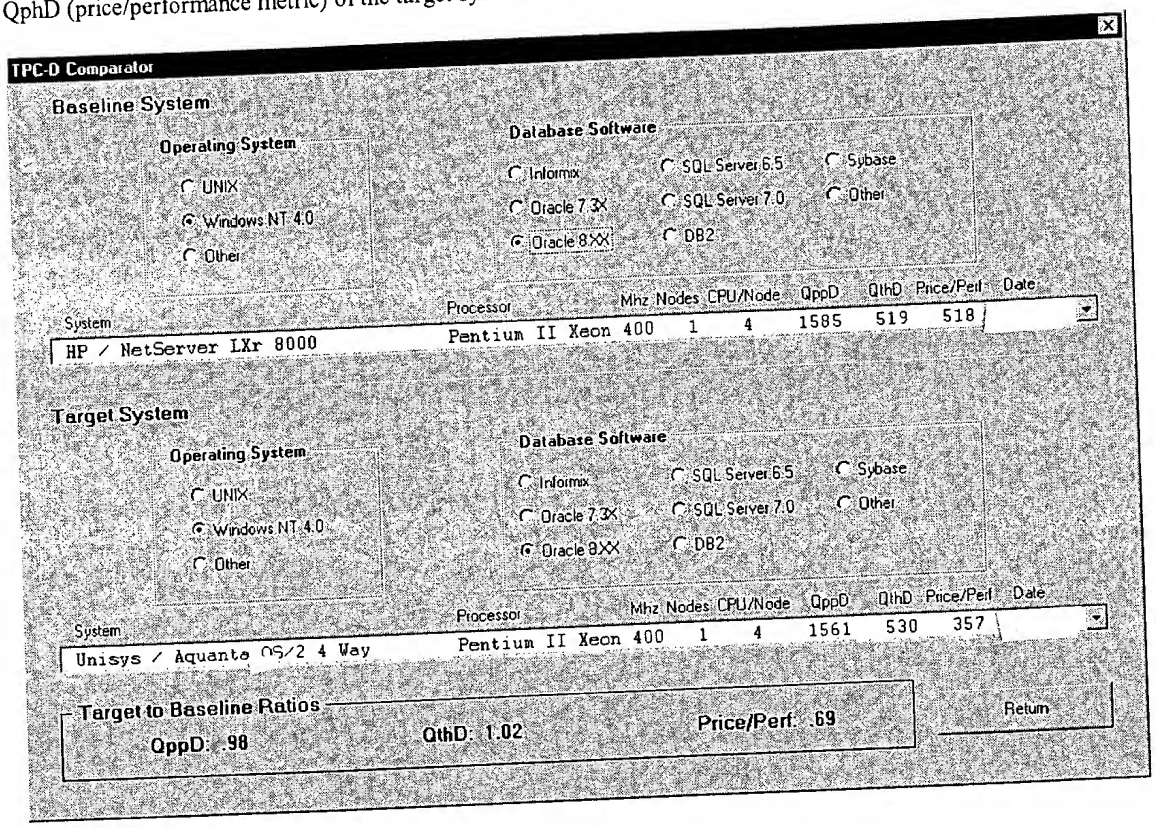
The comparison is completed by repeating the above process for the target system. When this is done, the ratios of tpmC and $/tpmC are shown at the bottom of the dialog.

We note also that the Comparator responds in a dynamic manner to each change or choice by the user. Thus, no "re-calculation" request is required by the user for each change made by the user.

In the example shown above, TPC-C value for the Unisys Aquanta QS/2 system, configured with four (4) 400 Mhz Intel Xeon processors is compared to that of an Acer system, configured with four 200MHz Pentium Pro processors. The tpmC Ratio shows that the Unisys system is yields a 60% higher throughput rate at 7% less cost.

3.2.2 TPC-D Comparator

To start the TPC-D Comparator, select **Comparator**, **TPCD Comparator** from the **Sizer Menu**. A dialog box, like the one shown below, will be displayed where you can choose a number of values for a baseline and target system. The Comparator will then calculate the ratio of QppD's (power metric), QthD (throughput metric), and $ / QphD (price/performance metric) of the target system to the baseline system.



The power metric, denoted QppD@Size, is calculated as the product of the relative database size and the reciprocal of the geometric mean of the timing intervals. The relative database size is the scale factor used to determine the population of each table in the database. For example, a scale factor of 1 yields about 1 GB of raw data, etc. Thus,

$$QppD@Size = 3600 * SF / [(Q_1 * \ldots * Q_{17}) * UF_1 * UF_2] ^ (1/19)$$

where

Q_i = Elapsed time to run query i within a single query stream
UF_1 = Elapsed time to run update function 1
UF_2 = Elapsed time to run update function 2
SF = Scaling Factor

Elapsed times are in seconds. The units of the power metric are queries per hour times the scale factor, thus the 3600 multiplier.

The throughput metric, denoted QthD@Size, is calculated as the ratio of the number of queries executed to the length of the measurement interval, i.e.,

QthD@Size = (no. of query streams) * 17 * 3600 * SF / (Length of measurement interval)

Length of measurement interval is in seconds. The units of the throughput metric are queries per hour times the scale factor.

The price/performance metric, denoted as $/QphD@Size, is calculated as the ratio of the total system price to the composite query per hour rating which is the geometric mean of QppD@Size and QthD@Size, i.e.,

$/Qphd@Size = (total system price) / (QppD@Size * QthD@Size) ^ (1/2)

The units of the price/performance metric are dollars per (queries per hour times the scale factor).

The operation of the TPC-D Comparator is identical to that of the TPC-C Comparator. Of course the exceptions are in the specific results. In this case we obtain the ratios of the power metrics, the throughput metric, and the price/performance metric.

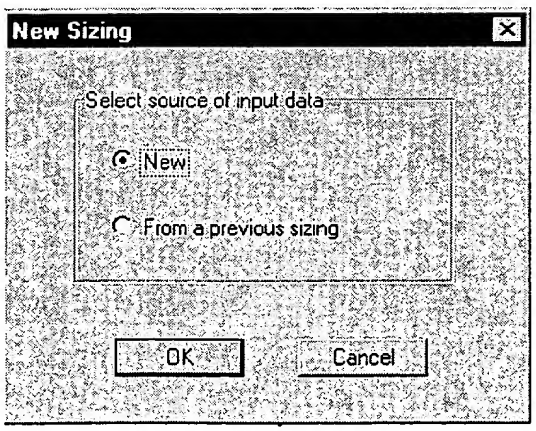
In the example shown above, TPC-D value for the Unisys Aquanta QS/2 system, configured with four (4) 400 Mhz Intel Xeon processors is compared to that of an HP system, configured with four (4) 400MHz Intel Xeon processors. configurations. The QppD andQthD ratios show that the Unisys system is within 2% of the HP system; however, the Uniys system costs approximately 30% less.

3.3 Mass Storage Sizing

A user may wish to determine only mass storage capacity requirements initially. This capability is provided via the Mass Storage Sizer feature.

The calculations of mass storage requirements for databases are based on an analysis of the table and index structures for each of the DBMS supported by the sizer. The estimates of file size requirements to support these databases are based on both recommendations from the DBMS vendors as well as experience in the field.

To estimate mass storage requirements select **Mass Storage Sizing** from the **Sizer Menu.** The dialog shown below is then displayed.

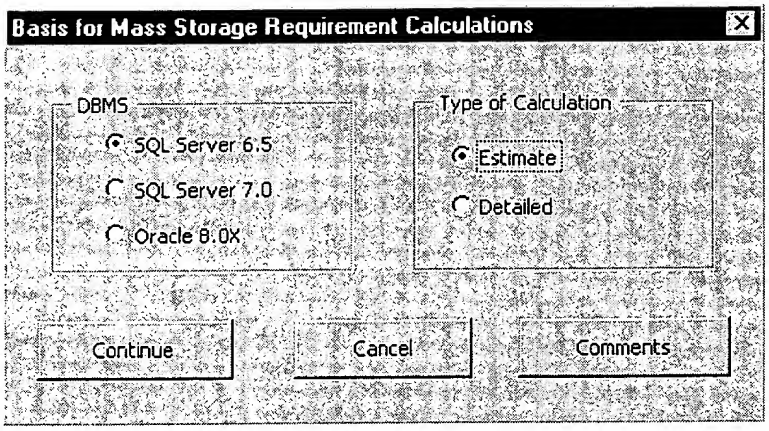


The user selects whether this is a new mass storage sizing or this is a carry over of a sizing that was previously saved to a workbook.

If the user selects **From a previous sizing**, an open file dialog is returned and the user can then select the file containing the results of a previous sizing. Note that the worksheet(s) in the previously saved workbook must have the same format as either of the two templates **Mass Storage Estimate.xlt** or **Mass Storage Estimate Detailed.xlt**. Upon selecting the file, the appropriate dialog is opened to modify the existing database definition, and make subsequent estimates.

If the user selects **New**, the following dialog, **Basis for Mass Storage Requirement Calculations**, prompts the user to indicate first, for which DBMS the sizing estimates will be made, and second, the type of calculation, based on the availability of information about the database. The Sizer currently supports estimation of mass storage requirements for SQL Server 6.5, SQL Server 7.0 and Oracle 8.04 databases. Also, estimates of mass storage requirements can be made, based on two levels of customer knowledge of the database, i.e.,

- **Estimate**: Little is known about the database
- **Detailed:** Aggregate sizes of rows and indexes are available by table



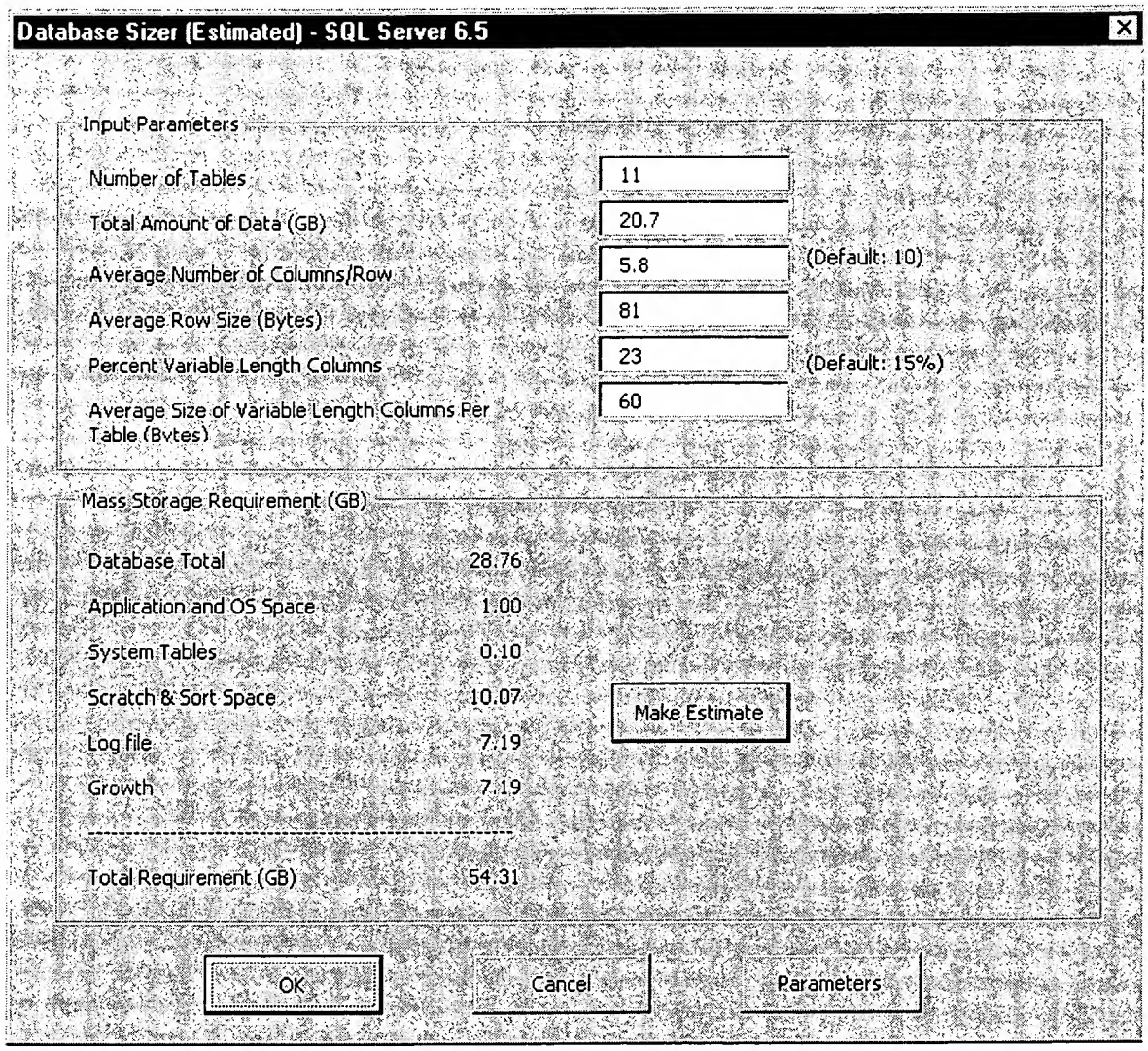
Click **Continue** to proceed to the next dialog window. In the following two sections the dialogs for the two levels of calculations, **Estimate** and **Detailed**, are described. The example database used is based on the Pubs Database that ships with SQL Server. The examples used are for the SQL Server 6.5 DBMS.

3.3.1 Estimate Calculation

The **estimated** case is based on the user providing six items of information, shown in the following table.

Input Item	Description
Number of Tables	Total number of tables in the database
Total Amount of Data, GB	Total estimated size of the raw database, in gigabytes
Average Number of Columns/Row	The average number of columns per row for all tables in the database
Average Row Size, Bytes	The average size of a row in a database table, expressed in bytes.
Percent Variable Length Columns	The percentage of columns for all tables that are variable in length, e.g., *varchar*. Defaults to 15%.
Average Size of Var. Length Columns per Table, Bytes	The aggregate number of bytes for all variable length columns in a table, averaged over all the tables

The dialog used to make the estimates is shown below.



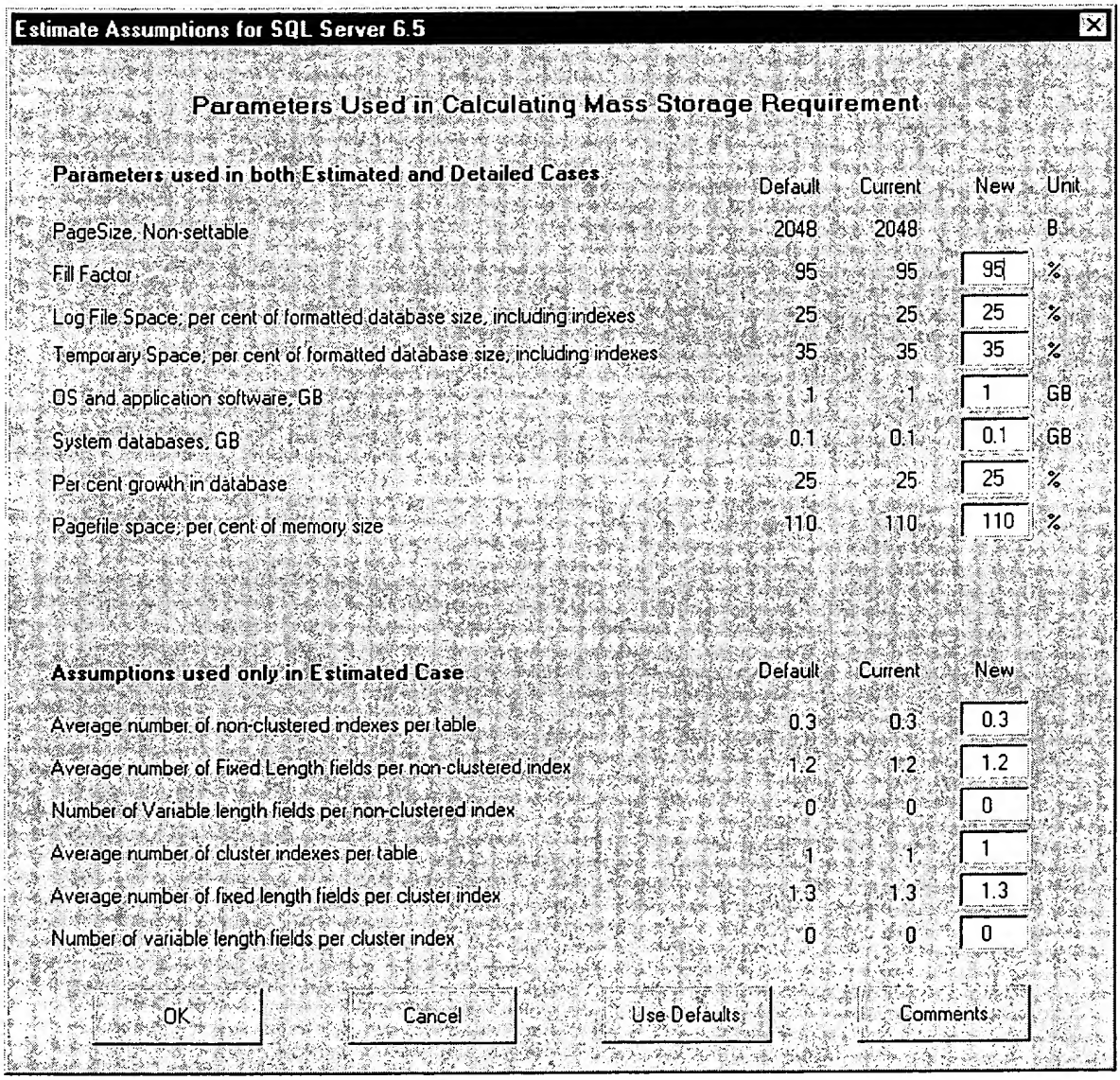
After having entered the six parameters characterizing the database, the mass storage requirement estimate is made and displayed via a click of the **Make Estimate** button. We note that the size of the page file is not included in these estimates. This is because the pagefile size is a function of the memory size requirement which is a function of the application load and CPU requirement.

Selecting **OK** or **Cancel** causes the form to be unloaded. Selecting **OK** gives the user the choice of saving this information to a workbook having the same format as the template **Mass Storage Estimate.xlt**. For a subsequent mass storage sizing, this same workbook can be loaded, the information contained in it is loaded into the sizer database and corresponding form. Note also that selecting **OK** keeps the sizing information in the sizer database so that in a subsequent sizing, during the same sizer session, this information is loaded into the form. Selecting **Cancel** causes the sizing information to be cleared.

In order to use the few parameters listed above to make the mass storage requirement estimates, certain additional assumptions are required in the areas of indexing and other space requirements. The indexing assumptions are

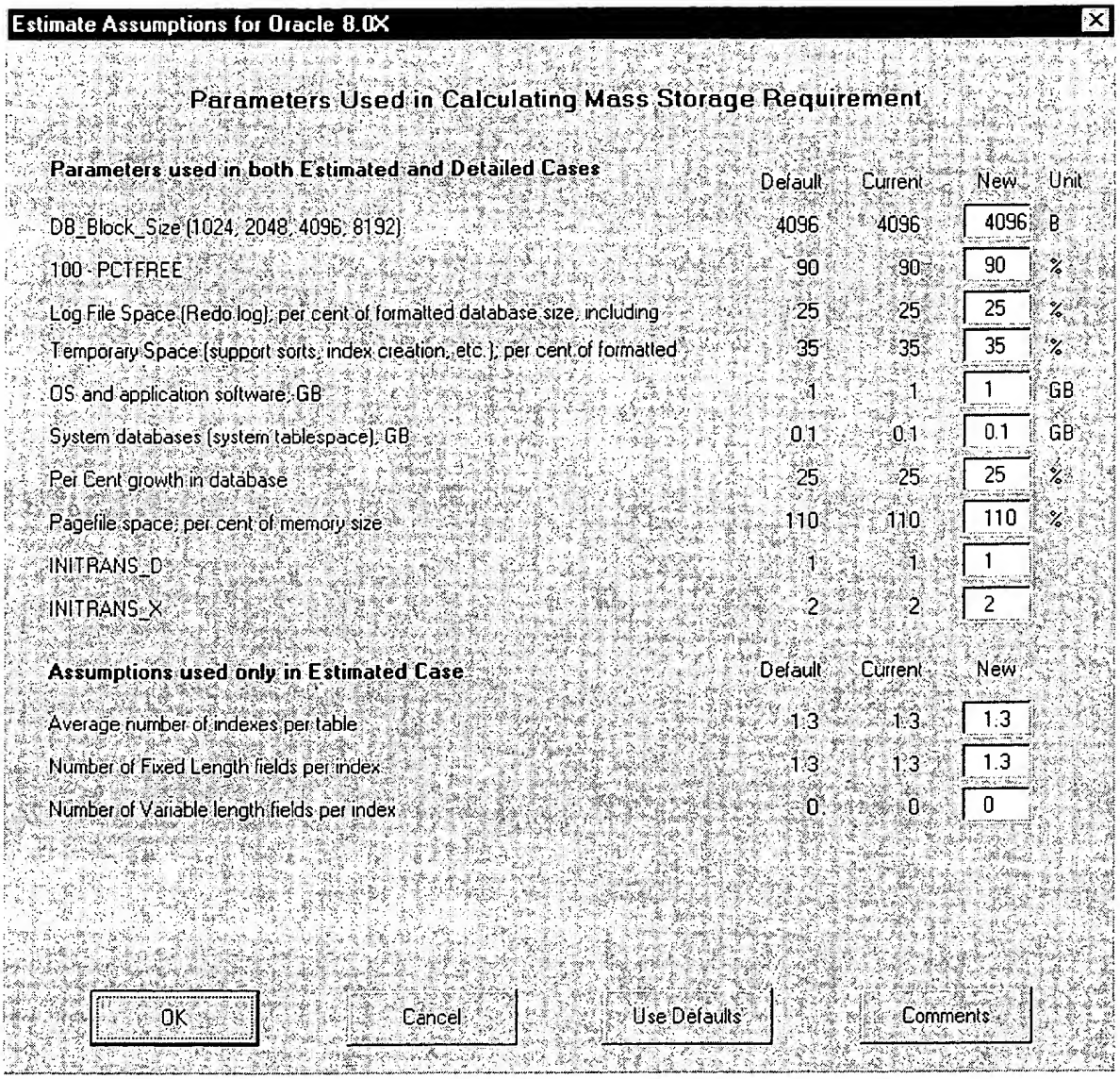
based on experiences in the field with various client databases. The space assumptions when using the SQL Server DBMS are based on Microsoft's recommendations for SQL Server 6.5 and Windows NT 4.0. The space assumptions when using the Oracle DBMS are based on Oracle recommendations as well as Unisys experience with its proprietary RDBMS.

A dialog is provided to show not only the default parameters values, but also to allow the user to change the values to whatever may be more applicable to the application being considered. This dialog, which is accessed via the **Parameters** button, is shown below. Note that the rightmost column contains values that can be modified. Also, the user can easily revert to the default values via the **Use Defaults** button. Further information on parameters is given via the **Comments** button.



The above dialog has the same appearance for each sizer supported DBMS. The set of default values for SQL Server 6.5 and SQL Server 7.0 is the same with the exception of the page size which is not modifiable (2048 for SQL Server 6.5 vs. 8192 for SQL Server 7.0). For Oracle, the parameters are as shown in the corresponding, following dialog. Note that Oracle's indexes using a B-Tree always have a leaf page level. Consequently, Oracle

does not have an equivalent "cluster" key, as defined for SQL Server. Addtionally, note that the sizer supports only size estimates for indexes characterized via B-Tree algorithms. Thus, the calculation of size requirements for cases of Oracle's "clustered tables and indexes" and "hash indexes" are not supported by the sizer.



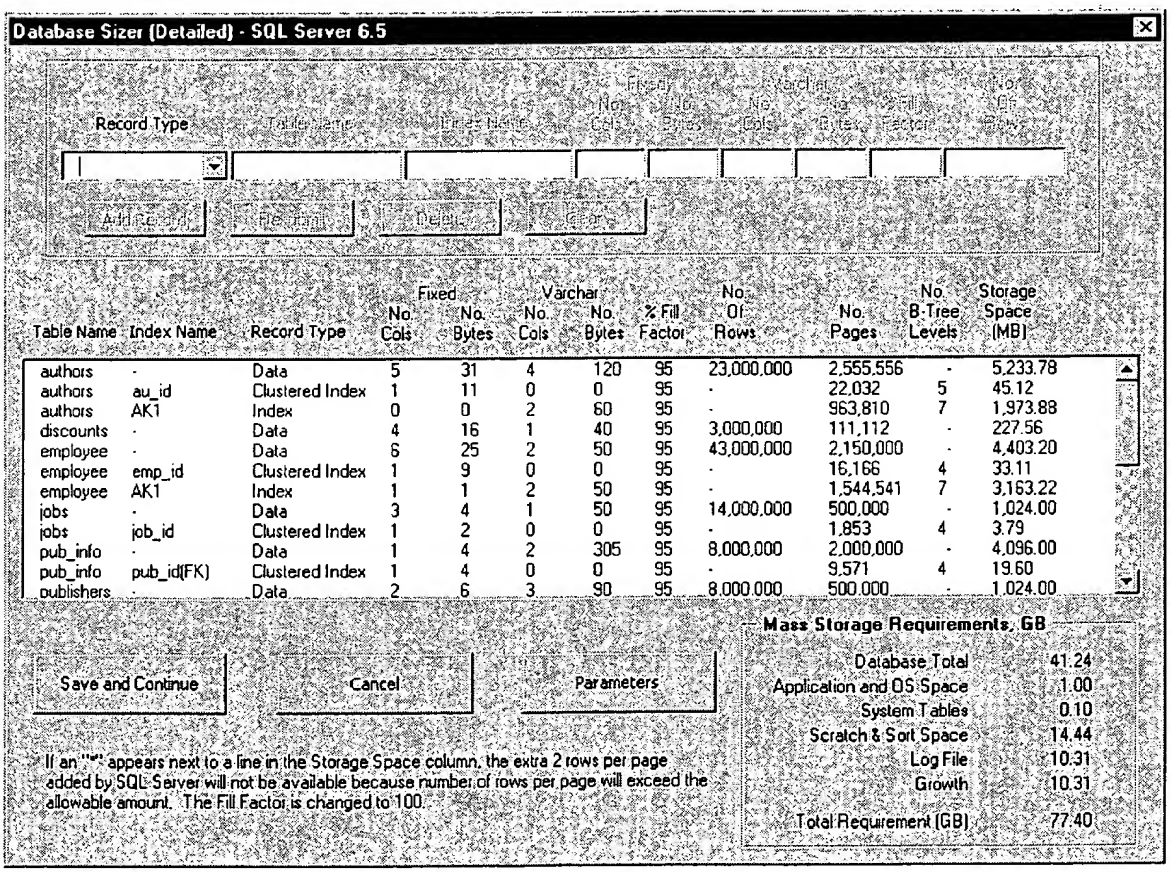
3.3.2 Detailed Calculation

The **Detailed** case is based on the user providing information about the size of the rows in each table and index in the database as well as the amount of space that should be made available for new rows or updating rows. This information includes the following:

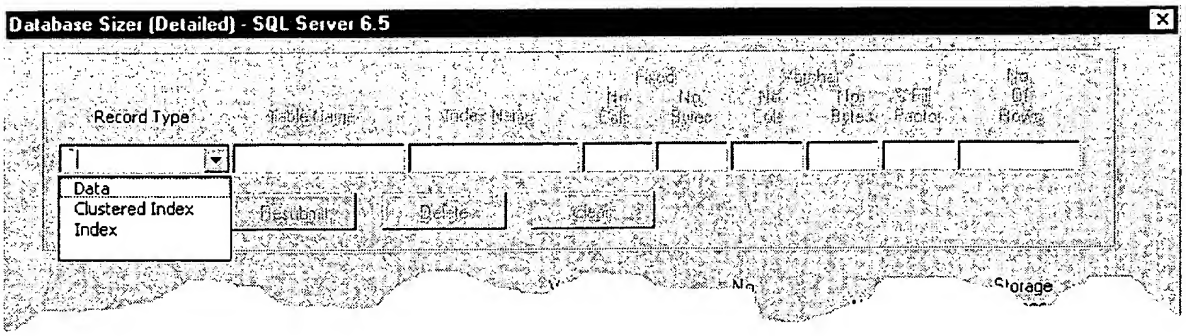
- For each table
  - Number of fixed size columns
  - Total fixed bytes per row
  - Number of varchar columns

  - Total varchar bytes per row
  - Number of rows
  - Fill Factor (SQL Server) or 100 - PCTFREE (Oracle)
- For each index (or clustered index for SQL Server case)
  - Number of fixed size index columns
  - Total fixed bytes of index columns per row
  - Number of varchar index columns
  - Total varchar bytes of index columns per row
  - Fill Factor (SQL Server) or 100 - PCTFREE (Oracle)

The dialog used to provide this data to perform the calculations is shown below. If the user had selected **New** sizing and a previous, detailed had been completed during the same sizer session, or, if the user had selected **Previous Sizing** and loaded a workbook with detailed sizng results, then the dialog would be loaded with the description of the database as is shown below. Otherwise, the form would be empty.

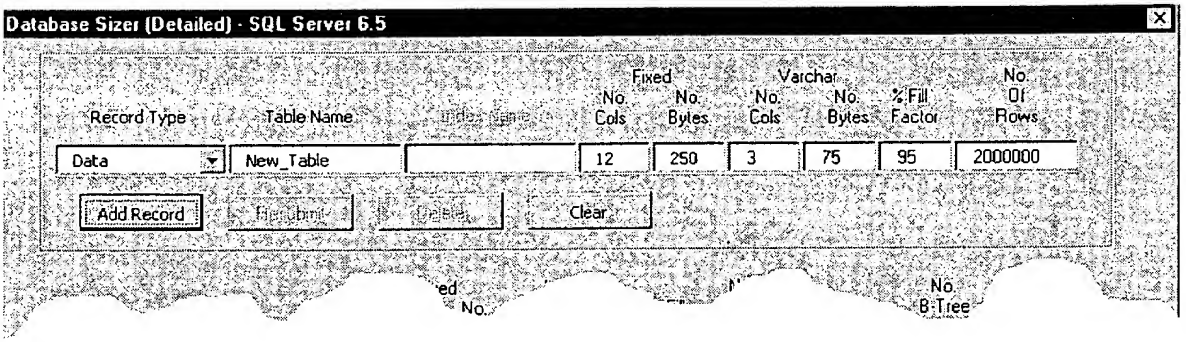


All input and modifications are made in the upper text boxes and upper buttons of the form. The user first selects the type of entry from the **Record Type** drop down, as shown in the next dialog. **Data** is selected if the user is going to add the characteristics for a table. Similarly, **Clustered Index** or **Index** is selected if the user is going to describe the characteristics for an index of a previously defined table.

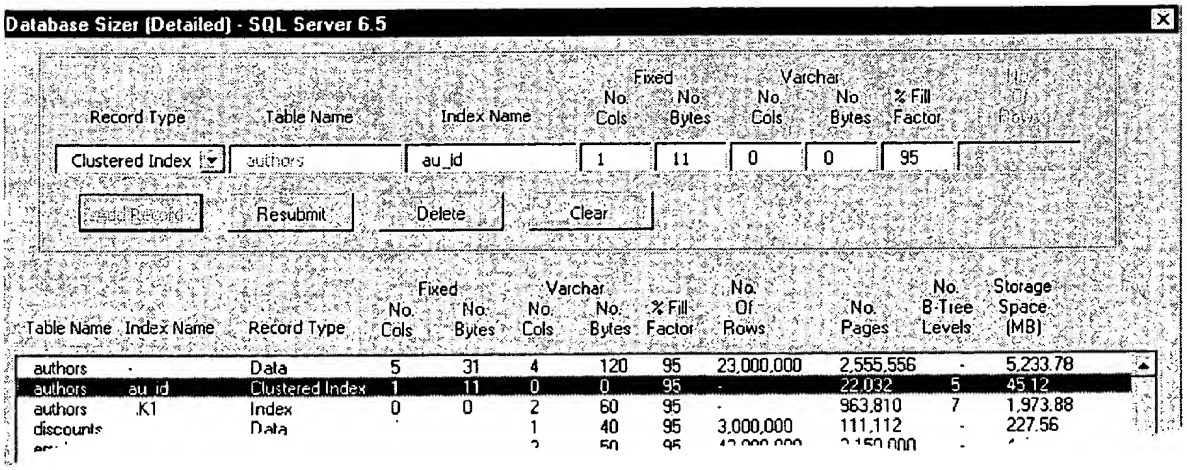


When a **Record Type** selection has been made, the text boxes and the **Add Record** button are enabled, as shown in the next dialog. The user then supplies the information in each text box. Once the data has been supplied, clicking the **Add Record** button will cause this entry to be added to the list box and the appropriate calculations to be made.

Note that when there are more than one table defined, the **Table Name** textbox becomes a drop down combo box listing all of the tables. This is useful when adding an index to a table.



Modfications can also be made to existing entries. This is done by first selecting the entry in the list box, as shown below.



This causes the entry to be displayed in the first line of the enabled text boxes. After modifications are made as necessary in the text boxes, the user clicks **Resubmit** to make the list box entry change and to perform the

necessary recalculations. Selecting **Delete** will cause the list box entry to be deleted; note also that if you are deleting the information for a table, then the corresponding index information is also deleted. Selecting **Clear** will cause the text boxes to clear and to de-select the list box entry.

The output from each entry provides the following information:

- The mass storage requirement for each table and each index
- The number of B-Tree levels associated with each index. This is the number of pages (SQL Server) or blocks (Oracle) that must be read from the disk and/or memory (cache) in order to access the first item of data.

Additionally, the box in the lower right corner provides summary information about the mass storage requirements. Note that this summary information is in the same format as that for the **Estimated** case.

Selecting the **Parameters** button will open the same dialog as the one shown for the **Estimated** case. Selecting **Save and Continue** or **Cancel** will unload the form. If **Cancel** is selected, the information in the form is cleared. If **Save and Continue** is selected, the user is then solicited to save the results to a workbook having the same format as the template **Mass Storage Estimate Detailed.xlt**.

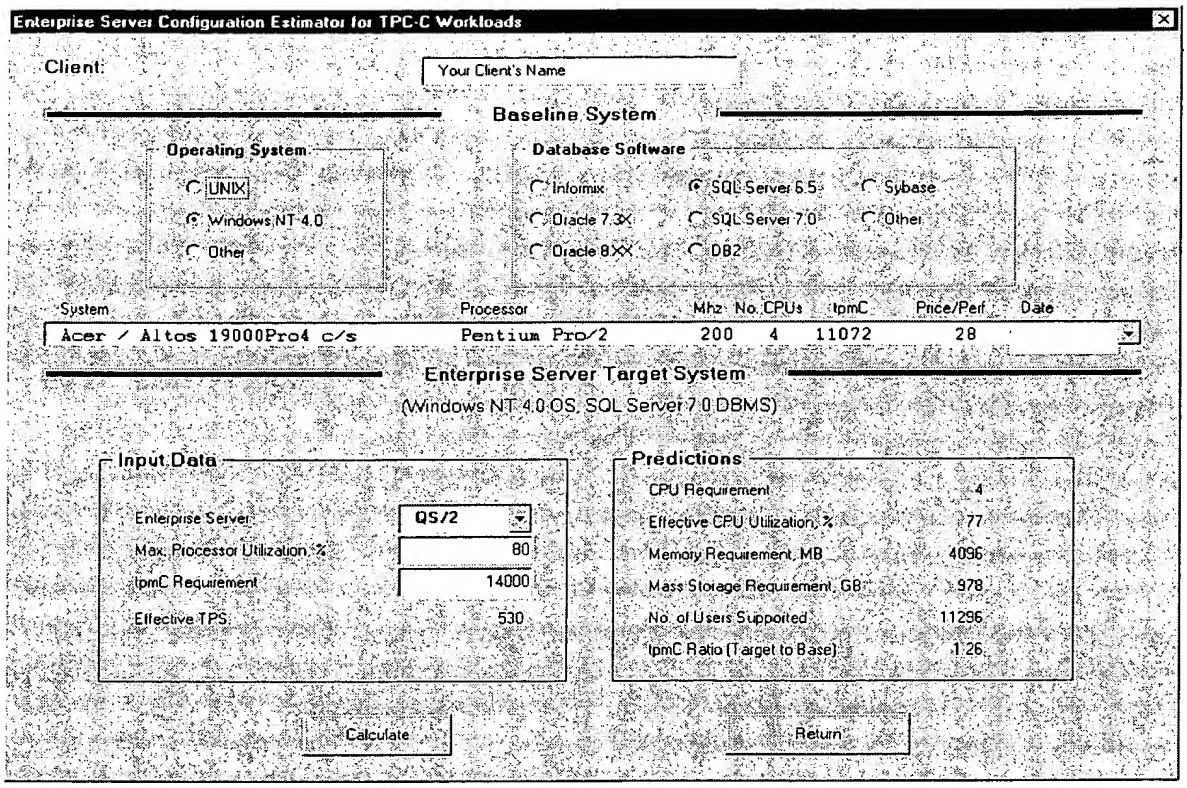
Note also that selecting **Save and Continue** keeps the sizing information in the sizer database so that in a subsequent sizing, during the same sizer session, this information is loaded into the form. This capability allows the user to switch from a detailed sizing for one DBMS to a detailed or estimated sizing for the same or different DBMS. Similarly, the user can switch from an estimated sizing to an estimated sizing for the same or different DBMS; however switching from estimated to detailed with the same information is not possible.

3.4 Using the NT Sizer

Currently, the Enterprise Sizer for Windows NT Applications estimates configuration requirements for OLTP workloads. A TPC-C workload may be selected, or the user may define his own. For the TPC-C workloads, one may also compare the resulting tpmC estimates to those of existing systems.

3.4.1 TPC-C Workloads

To size a TPC-C workload, select **OLTP System Sizing, TPCC Workload** from the **Sizer Menu.** The dialog shown below is then displayed.



The top half of the dialog provides the information from the TPC-C Comparator to allow you a basis or baseline to which you can compare your projected target system estimates. The lower left quadrant of the dialog allows you to select the target system and to specify its performance characteristics.

Estimated and measured TPC-C performance data is available for NT servers running the SQL Server 7.0 DBMS on the more recent and future systems; these systems include the XR/6 with up to 12 processors, the QS/2 with up to 4 Xeon processors, and the to-be-announced QS/2 follow-on (FO) which will be offered with configurations up to 8 processors. The "target" systems for sizing purposes comprise these systems. Older systems are not included in the set of target systems.

For the target system to be sized, you must specify the following as indicated in the lower left quadrant of the dialog:

- The specific server from the choices of XR/6, QS/2, and QS/2 FO
- Maximum processor utilization
- tpmC requirement

The maximum processor utilization is the processor utilization level that you do not wish to exceed with the specified workload on the proposed system. Specifying a maximum of 100% is not recommended, as response times degrade as the processor utilization approaches 100%, and it also provides no room for growth. Specifying a value too low will provide configuration requirements that far exceed the input requirements. If you do not have a processor utilization number in mind, use 80-85% as this will provide a reasonable estimate with a safety margin.

Since this type of sizing is based on a TPC-C workload, you must also specify the tpmC requirement. If you do not know what tpmC value to use, start with a baseline system and increase or decrease the tpmC value accordingly.

Once the baseline and target system characteristics have been selected, click the **Calculate** button. Results of the calculations are shown in the lower right quadrant of the dialog.

For example, in the dialog above, the Unisys QS/2 server with the Xeon 400 Mhz processor is being compared to the Acer 4 way system with the 200 Mhz processor. The tpmC requirement was input as 14000 on the QS/2 server with a requirement that the processor utilization not exceed 80%. This results in a configuration requiring 4 processors that will operate at 77% utilization on the average. Further, this system provides 26% more throughput than the baseline. Also shown are the estimated processor and memory requirements of 4096 MB and 978 GB, respectively.

Note that when you enter or change a tpmC value, the corresponding value for **Effective TPS** changes automatically. This number represents the total number of *transactions per second* that could be realized for the target system with the specified tpmC value. The tpmC value represents *transactions per minute* for TPC-C **New Order** transactions, where New Order transactions represent approximately 45% of the total transaction workload.

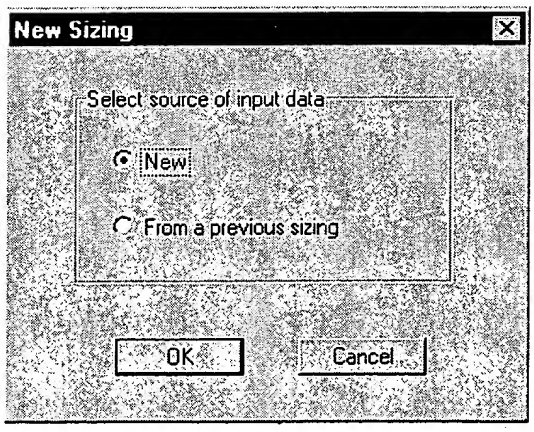
3.4.2 User Defined Workloads

This section will show you how to use the Sizer to estimate system configuration requirements for a user-defined OLTP workload. It will guide you through the steps required to define the database and the application workload characteristics.

3.4.2.1 Previous vs. New Sizing

Select **OLTP System Sizing, User Defined Workload** from the **Sizer Menu** to size a user-defined OLTP workload. You are then presented with the **New Sizing** options window shown below. You have the option of performing a totally new sizing or continuing from a sizing which was previously saved to an Excel workbook. Whether you select **New** or **From a previous sizing**, you will be lead through a three step process which will produce a system configuration estimate.

If you select a previous sizing, you will be queried for the workbook name via the standard Excel **Open File** dialog. We note that if you select a file not previously saved via a sizing, or altered since it was saved as sizing results, the sizer does not necessarily recognize this; consequently, the results are unpredictable. The workbook selected must have the same formatted worksheets as those in the template UserBk.xlt.



3.4.2.2 Hardware Selection

The main dialog from which the multi-step sizing process functions is shown below. One first selects the system type and the desired backbone LAN speed. The various systems from which to select are the following:

- XR/6 with 200 Mhz Pentium Pro processor
- QS/2 with 400 Mhz Xeon processor
- QS/2 with 450 Mhz Xeon processor
- QS/2 follow-on (FO) with 400 Mhz Xeon processor
- QS/2 FO with 450 Mhz processor

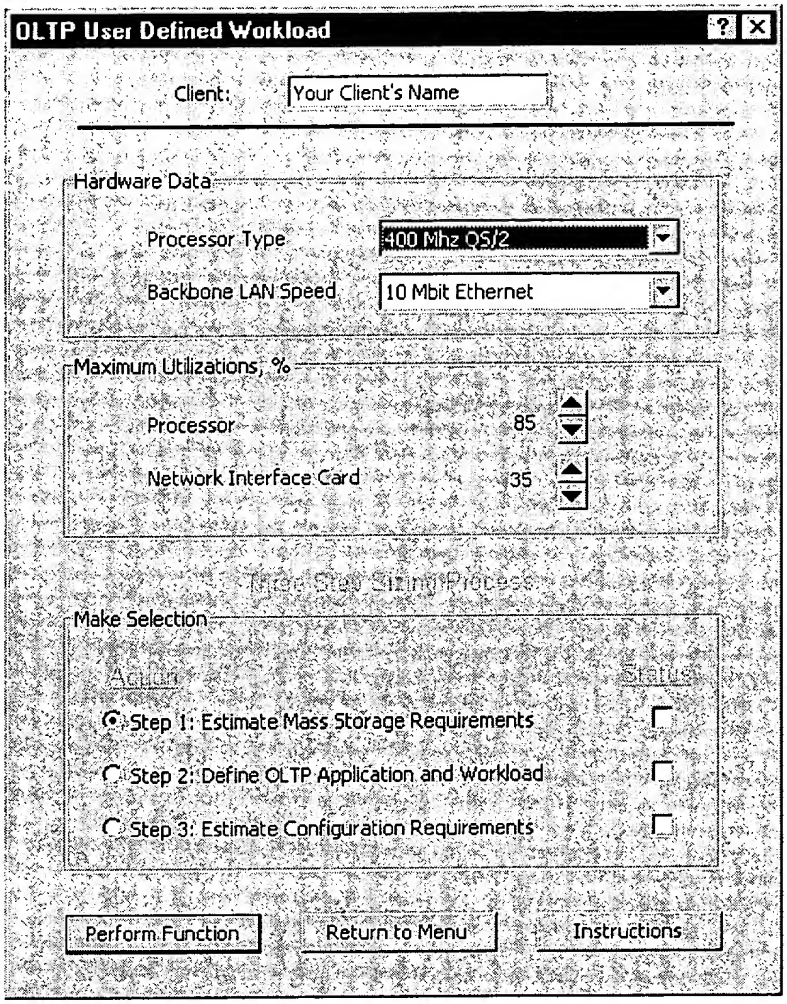
A corresponding maximum processor utilization is also specified.

We note that some of these systems are not as yet available; however, there is available data from which we can estimate system requirements. Thus, this provides some predictive capability for our future systems.

The various LAN speeds from which we can select are the following:

- 10 Mbit Ethernet
- 10 Mbit Switched Ethernet
- 100 Mbit Ethernet
- 100 Mbit Switched Ethernet
- 1 Gbit Switched Ethernet
- Best Fit

Selecting Best Fit allows the Sizer to determine the smallest LAN speed that will satisfy the expected LAN traffic at or below an optimal, maximum utilization. The optimal utilization is dynamically entered and displayed with the Network Interface Card maximum utilization whenever the corresponding LAN speed is selected. Additionally, this value can be overridden by the user. For example, selecting a 100 Mbit LAN speed, the maximum, optimal utilization is considered to be about 35% which is entered as the Network Interface Card utilization; the user can then optionally override this value using the corresponding spinner.



Having selected the hardware, the next two steps are to estimate mass storage requirements and define the application and workload. Selecting the hardware and these two steps can be done in any order. To estimate mass storage requirements, select the radio button corresponding to **Step 1**, and click the **Perform Function** button.

3.4.2.3 Estimate Mass Storage Requirements

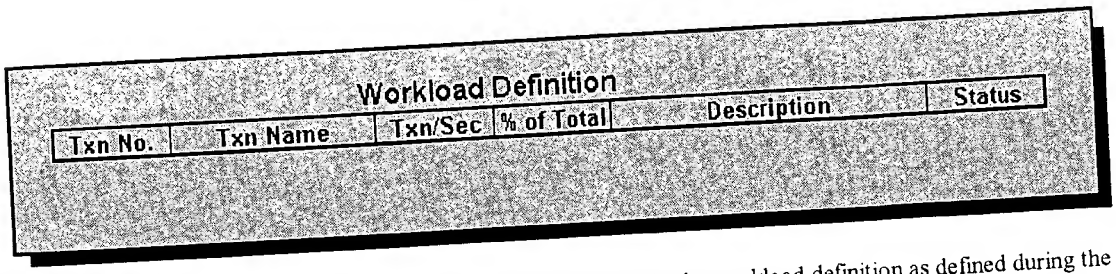
The process to estimate mass storage requirements was described in Section 3.3. The process described in that section is identical to this portion of the system sizing process.

Selecting Step 2, we can now define the application and its workload.

3.4.2.4 Define OLTP Application and Workload

The user is taken to a worksheet from which he performs a series of steps to define the application and the workload. A transaction consists of a series of SQL statements surrounded by a BEGIN TRANSACTION and a COMMIT.

If you are conducting a new sizing, you will be presented with a blank worksheet that looks like the one shown below. You must use the **Workload** menu options to define the number and content of the transactions. This process is described in the following sections.



If you are using a previous sizing, the worksheet will already contain the workload definition as defined during the last sizing.

3.4.2.4.1 Workload Definition Process

A worksheet, shown below, is processed using the **Workload** menu which consists of the following five items used to define the application and the workload.

Menu Item	Description
Instructions	Provides a set of online instructions to guide you through the workload definition phase.
Add Transactions	Adds the specified number of transactions to the application.
Delete Transactions	Deletes the specified transaction from the application.
Transaction Composition	Allows you to define SQL statements for a specified transaction.
Return	Return to the three step sizing process main menu.

A recommended sequence is the following:

- **Specify how many unique transactions comprise the application:** From the "Workload" menu, selecting "Add Transactions" will allow you to add a user supplied number of transactions to the application. The added transactions are highlighted via shading.
- **Generically identify each transaction by name:** Fill in the transaction name column, and provide a short description, if desired.
- **Specify each transaction's load on the system:** Fill in the expected rate in the transaction per second (TPS) column.
- **Specify each transaction's composition: SQL inserts, deletes, updates, selects:** This is done via a subsequent dialog as a result of selecting the Transaction composition button. This process is described in detail in Section 3.4.2.4.2.

Workload Definition

Txn No.	Txn Name	Txn/Sec	% of Total	Description	Status
1	New Order	35	26.9%		Defined
2	Payment	35	26.9%		Defined
3	Add Publisher	10	7.7%		Defined
4	Delivery	35	26.9%		Defined
5	Delete Stores	15	11.5%		Defined
Totals		130	100.0%		

This process is repeated until all of the transactions have been defined. Note that you can both add and delete transactions during this phase, thus allowing you to further modify the application and workload definition. Select

the "Instructions" menuitem from the Workload menu for extra help on completing the information in this worksheet.

3.4.2.4.2 Composition of OLTP Transaction

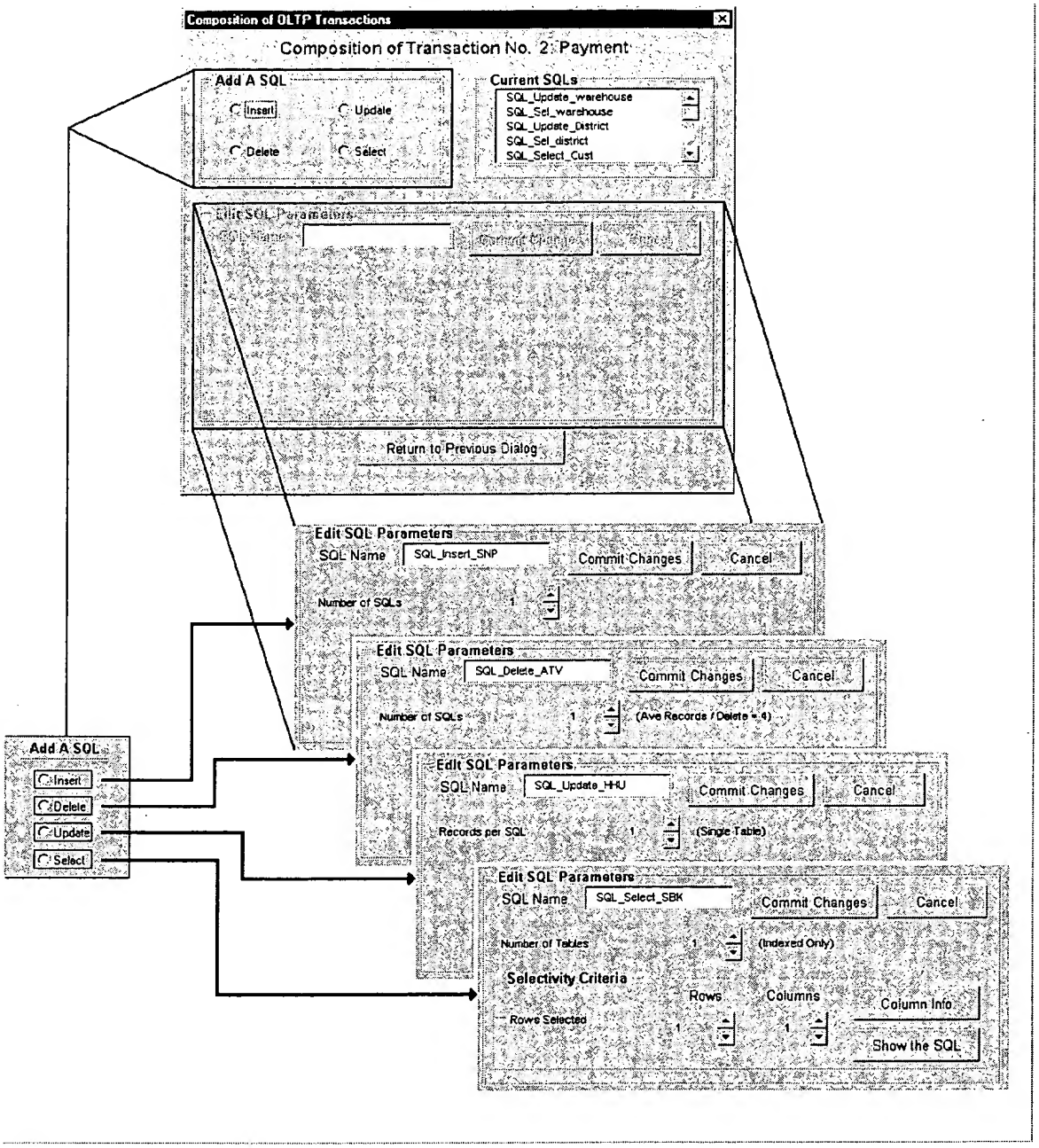
Recall that each transaction can consist of several SQLs. To activate the transaction composition dialog shown below, highlight a transaction in the worksheet and then select the **Transaction Composition** menuitem from the **Workload** menu. In this dialog we can now specify the various types of SQLs comprising a given transaction. We can add an Insert, Delete, Update, or Select SQL by selecting a radio button in the **Add a SQL** group. Selecting one of the four SQL types causes a corresponding menu to be displayed in the **Edit SQL Parameters** group located in the lower half of the dialog; each of these menus is also shown in the dialog below. A default **SQL Name** is also entered with its suffix randomly generated to guarantee uniqueness. You can use the generated name or highlight it and enter your own name for the selected SQL statement.

Choosing an Insert, Delete, or Update command will highlight the corresponding **Number of SQLs** item. Use the spinner to specify the number of Insert or Delete SQLs, or the number of records Updated.

For the Select statement, use the highlighted spinner to specify whether it is a single table Select, or a nested join from either two or three tables. In all cases, the assumption is that the Selects are indexed, which is consistent with OLTP applications. For all Select cases, the **Selectivity Criteria** must also be completed. For a Select from a single table, the selectivity refers to the number of rows selected from the table. For a two-table join, the selectivity refers to the number of rows selected from the outer, or left, table; the inner table selectivity is assumed to be an average of four for each row selected from the outer table. For a three-table join, the selectivity is equal to the number of values in the WHERE clause that pertain to the outer table. For further clarification on how each of the select cases may be applied in a specific OLTP application, click the corresponding **Show the SQL** button and refer also to section 4.

Additionally, for the Select statement, the number of columns and aggregates per row returned are requested. These values together with the number of rows returned are used to estimate the amount of traffic on the LAN per SQL and transaction. The number of columns per row are not currently used to calculate processor usage.

For each SQL added, you must save the changes by clicking the **Commit Changes** button. This action causes the newly created SQL to be added to the **Current SQLs** list box. You may also modify your definition of each SQL by selecting the SQL from the list box. This causes its definition to be highlighted in the **Edit SQL** group. Upon completion of editing, commit the changes.

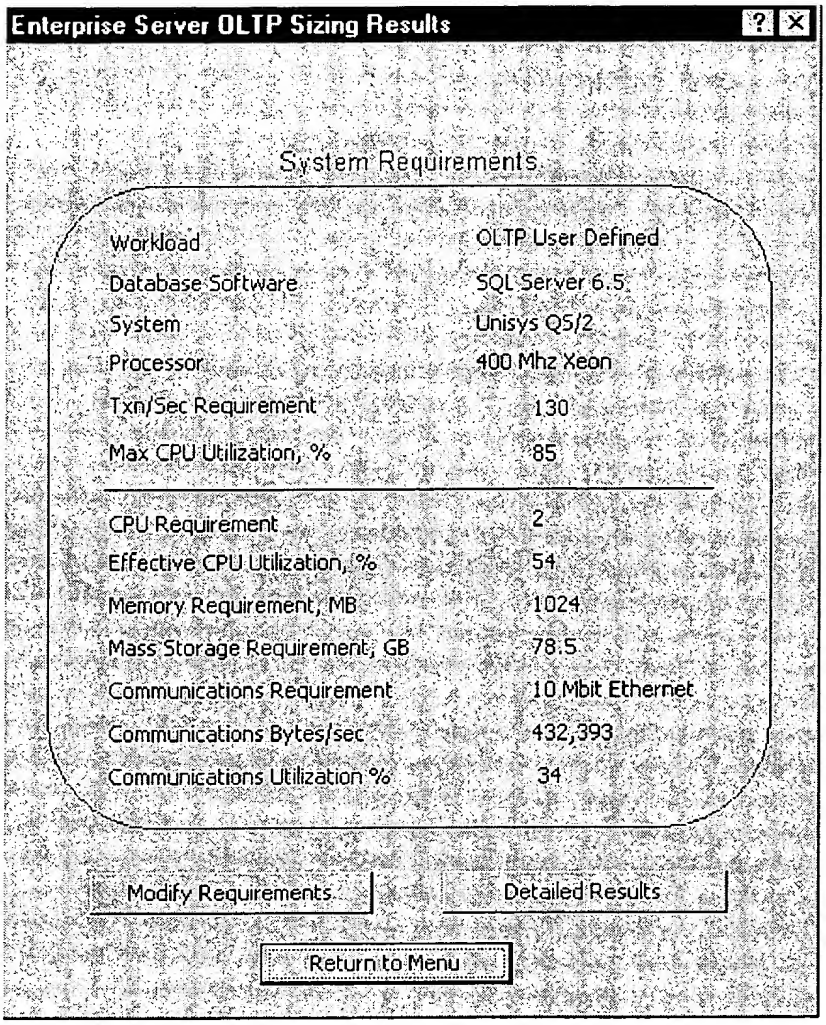


Once you have defined this transaction, select **Return to Previous Dialog**. This takes you back to the worksheet where you can select another transaction for further definition. As noted previously, you can always return to this dialog for further definition of a selected transaction. Continue this process until all transactions are defined.

3.4.2.5 Estimate Configuration Requirements

Once all of the transactions have been defined as in section 3.2.2.2, we now have enough data to estimate the configuration requirements. Select the radio button for **Step 3** and click **Perform Function**. Note that the **Status** box must be checked for both steps 1 and 2 in order to proceed with the third step.

The sizer now estimates the configuration requirement. An example of the format of the results is shown in the following dialog. Input requirements are shown above the horizontal bar. The items below the bar are calculated by the Sizer.



The **CPU Requirement** indicates the number of processors required to support the defined workload; the processors are of the type specified by the **System** field. The **Effective CPU Utilization** is the estimated average total processor utilization for the specified workload. The **Memory Requirement** is given in megabytes (MB). The **Mass Storage Requirement,** given in gigabytes (GB), is the total amount of disk space required to support the given application; this includes following:

- Formatted database size
- Windows NT operating system, the DBMS, and application files
- DBMS system tables
- Scratch and sort space
- Transaction log file
- Paging file
- Growth

The type of ethernet card required to support this workload is indicated as well as the expected amount of LAN traffic along with the expected ethernet card utilization.

The buttons at the bottom of the dialog provide you with several options. Click **Modify Requirements** to return to the sizing main menu, where you can make changes to the database size or transaction workload assumptions and then have the Sizer recalculate the system configuration requirements. The current results will not be saved.

3.4.2.6 Saving Results

The **Detailed Results** button will save the sizing results onto the corresponding worksheets used in the UserBk*n*.xls workbook. The user is then prompted to save the resulting workbook to disk. The names and contents of the resulting worksheets are listed in the table below. A sample of each of the results worksheets is provided in the remaining sections.

Worksheet Name	Contents
Mass Storage Sizing – Summary	Summarizes the mass storage sizing. Includes input parameters on a database basis, summary output data, and the additional parameters used to make the calculations
Mass Storage Sizing - Details	If a detailed sizing was done, then this worksheet contains the per table details entered via the mass storage dialog.
OLTP Input - Workload	Contains the defined transaction workload.
OLTP Input - Transactions	Contains the defined transaction composition for each transaction.
OLTP Sizing Results	Overall results of the sizing.
OLTP CPU Load	Bar chart depicting relative CPU utilization for each transaction in the mix.
Comm Load	Bar chart showing relative usage of the ethernet interface card for each transaction
Capacities – CPU	Chart showing estimated peak transaction rates for multi-processor configurations.

The **Return to Menu** button will return you to the Sizer "home page". If you press this button without having exercised the **Detailed Results** option, the Sizer will not retain the sizing just completed.

3.4.2.6.1 Worksheet Results: Mass Storage Sizing

The estimated mass storage requirements are placed in one or two worksheets, depending on the type of mass storage sizing conducted. If the mass storage sizing was of the **Estimated** type, then the input parameters and results are placed in a worksheet called **Mass Storage Sizing – Summary**. If the mass storage sizing was of the **Detailed** type, then the input parameters and results are placed in both the **Mass Storage Sizing – Summary** and **Mass Storage Sizing – Details** worksheets.

For both the **Estimated** and **Detailed** types, the **Mass Storage Sizing – Summary** worksheet has the format shown in the "Mass Storage Sizing Summary" table shown below. The information includes the six input parameters specified for the **Estimated** case, the summary mass storage requirements, and the additional sets of parameters required to make the calculations. For a **Detailed** case the six input parameters are weighted averages of the data supplied to the detailed sizing. Note that the mass storage estimates exclude page file size. This is because the results contained here were calculated during the database sizing portion of the exercise, and paging file requirements are based on total memory requirements. Total memory size was not known until the final system configuration estimate was completed.

Mass Storage Sizing Summary

SQL Server 6.5

(Based on Detailed Database Information)

Summary of Input Data

Description	Value
Number of Tables	11
Total Amount of Data, GB	20.7
Average Number of Columns/Row	5.8
Average Row Size, Bytes	81
Percent Variable Length Columns	23
Average Size of Var. Length Columns, Bytes	60

Mass Storage Estimates

Description	Value
Database Size, GB	41.24
Application and OS Space, GB	1
System Tables, GB	0.1
Scratch and sort Space	14.44
Growth	10.31
Log Files	10.31
Total Mass Storage Requirement*	77.4

*Excludes page file

Assumptions Used in Calculations

Description	Value
PageSize, Non-settable	2048
Fill Factor	95
Log File Space, per cent of formatted database size, including indexes	25
Temporary Space, per cent of formatted database size, including indexes	35
OS and application software, GB	1
System databases, GB	0.1
Per cent growth in database	25
Pagefile space, per cent of memory size	110

For the **Detailed** case, the **Mass Storage Sizing – Details** worksheet has the format as shown in the "Database Statistics by Table" table below. This lists the same information as that given in the detailed sizing form.

Database Statistics by Table

Table Name	Index Name	Record Type	Fixed Columns	Fixed Bytes	VarChar Columns	VarChar Bytes	Fill Factor	Number of Rows	Number of Pages	B-Tree Levels	Storage Space (MB)
authors	-	Data	5	31	120	4	95	23,000,000	2,555,556	-	5,233.78
authors	au_id	Clustered Index	1	11	0	0	95		22,032	5	45.12
authors	AK1	Index	0	0	60	2	95		963,810	7	1,973.88
discounts	-	Data	4	16	40	1	95	3,000,000	111,112	-	227.56
employee	-	Data	6	25	50	2	95	43,000,000	2,150,000	-	4,403.20
employee	emp_id	Clustered Index	1	9	0	0	95		16,166	4	33.11
employee	AK1	Index	1	1	50	2	95	-	1,544,541	7	3,163.22
jobs	-	Data	3	4	50	1	95	14,000,000	500,000	-	1,024.00
jobs	job_id	Clustered Index	1	2	0	0	95		1,853	4	3.79
pub_info	-	Data	1	4	305	2	95	8,000,000	2,000,000	-	4,096.00
pub_info	pub_id(FK)	Clustered Index	1	4	0	0	95		9,571	4	19.6
publishers	-	Data	2	6	90	3	95	8,000,000	500,000	-	1,024.00
publishers	pub_id	Clustered Index	1	4	0	0	95		2,394	4	4.9
publishers	pub_name	Index	0	0	40	1	95		229,631	6	470.28
roysched	-	Data	4	18	0	0	95	86,000,000	924,732	-	1,893.85
sales	-	Data	4	20	32	2	95	21,000,000	724,138	-	1,483.03
sales	stor_id ord_num	Clustered Index	1	4	20	1	95		14,200	5	29.08
stores	-	Data	3	11	80	3	95	6,000,000	375,000	-	768
stores	stor_id1	Clustered Index	1	4	0	0	95	-	1,796	4	3.68
stores	stor_name	Index	0	0	40	1	95	-	172,223	6	352.71
titleauthor	-	Data	4	22	0	0	95	25,000,000	324,676	-	664.94
titleauthor	FK	Clustered Index	2	17	0	0	95		3,866	4	7.92
titles	-	Data	8	54	280	2	95	18,000,000	6,000,000	-	12,288.00
titles	title_id	Clustered Index	1	6	0	0	95	-	35,089	5	71.86
titles	title	Index	0	0	80	1	95	-	956,250	7	1,958.40

3.4.2.6.2 Worksheet Results: OLTP Input – Workload

The worksheet **OLTP Input - Workload** is created during Step 2 of the sizing process. It contains the specified transaction names and rates, and also a description for each one. The **Status** column on the right indicates that the transaction content was defined during the second step.

Workload Definition

Txn No.	Txn Name	Txn/Sec	% of Total	Description	Status
1	New Order	35	26.9%		Defined
2	Payment	35	26.9%		Defined
3	Add Publisher	10	7.7%		Defined
4	Delivery	35	26.9%		Defined
5	Delete Stores	15	11.5%		Defined
Totals		130	100.0%		

3.4.2.6.3 Results Worksheet: OLTP Input – Transactions

The worksheet **OLTP Input - Transactions** shows the composition of each transaction as defined during Step 2 of the sizing.

Transaction Composition

Txn No.	SQL No.	SQL Name	SQL Type	SQLs / Txn	Selectivity	Columns
1	1	2Sel	Sel-2 Tables	1	1	2
1	2	1Sel_District	Sel-1 Table	1	1	3
1	3	SQL_Update	Update	1	1	
1	4	SQL_Insert_Orders	Insert	1	1	
1	5	SQL_Insert	Insert	1	1	
1	6	SQL_Sel_1_item	Sel-1 Table	1	1	1
1	7	SQL_Sel_1_item	Sel-1 Table	1	1	1
1	8	SQL_Sel_1	Sel-1 Table	1	1	1
1	9	SQL_Sel_1_stock	Sel-1 Table	1	2	2
5	1	SQL_Select_distric	Sel-1 Table	1	1	1
5	2	SQL_Select_OL-sto	Sel-2 Tables	1	1	1

Legend:

SQL Type	SQLs/Txn	Selectivity
Insert	n	1
Delete	n	4
Update	1	n
Sel - 1 Table	1	n
Sel - 2 Tables	1	n
Sel - 3 Tables	1	n

3.4.2.6.4 Results Worksheet: OLTP Sizing Results

The worksheet **OLTP Sizing results** contains the same results as were displayed in the **System Requirements** report dialog. When the data is in worksheet format, it can be printed using the standard Excel printing options.

Enterprise Server Sizing Results for OLTP User Defined Workload

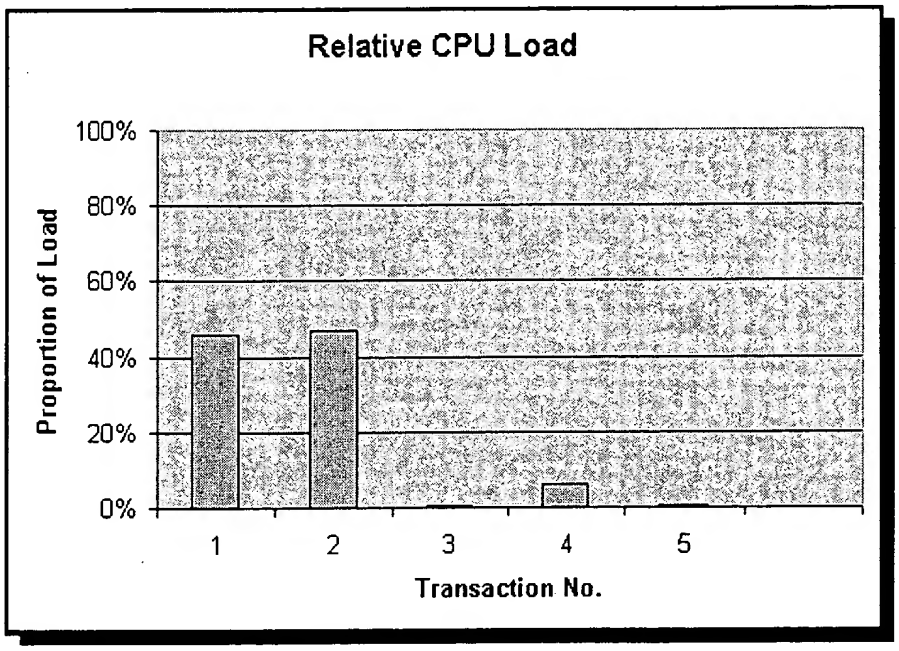
Client: Your Client's Name

System Requirements

Workload	OLTP User Defined
Database Software	SQL Server 6.5
System	Unisys QS/2
Processor	400 Mhz Xeon
Txn/Sec Requirement	100
Max. CPU Utilization, %	85
CPU Requirement	2
Effective CPU Utilization, %	46
Memory Requirement, MB	1024
Mass Storage Requirement, GB	78.5
Communications Requirement	10 Mbit Ethernet
Communications - Bytes/sec	393,798
Communications Utilization %	31

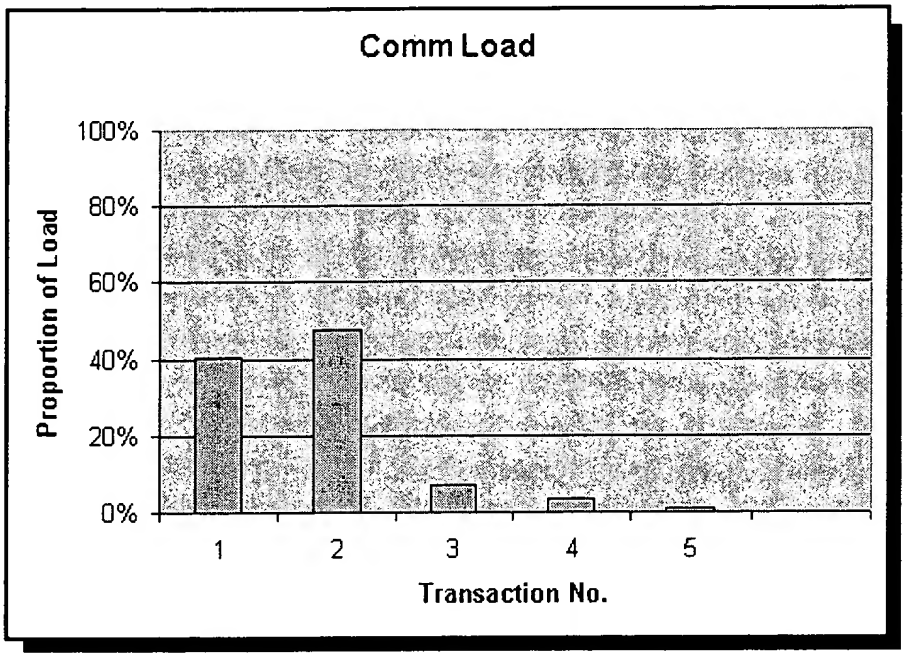
3.4.2.6.5 Results Worksheet: OLTP CPU Load

The bar chart in the **OLTP CPU Load** worksheet shows what proportion of the CPU load is attributed to each transaction. In the sample chart shown above, transaction 1 accounts for approximately 20% of the total workload, transaction 2 accounts for 70%, and transaction 3 is approximately 10%.



3.4.2.6.6 Results Worksheet: Comm Load

The bar chart in the **Comm Load** worksheet shows what proportion of the communications load is attributed to each transaction. In the sample chart shown above, over 90% of the load is attibuted to transaction 2.



3.4.2.6.7 Results Worksheet: Capacities - CPU

The **Capacities – CPU** worksheet contains a table showing the estimated peak transaction rate that can be achieved for each processor configuration based on the defined application workload.

The sample table shown below contains the estimates for the same sizing as used in previous examples. It shows that a maximum 243 transactions/sec could be achieved on a single processor system running the specified application workload. The original sizing estimate showed that the specified transaction rate of 31/sec could be achieved on a single processor system with an effective CPU utilization of 13%.

Estimated Enterprise Server Capacities

Client: Your Client's Name

Estimated Peak Transaction Rate (Transactions/Sec)

No. Processors	Capacity @ 100% CPU
1	113
2	219
3	320
4	416

System Data

Workload	OLTP User Defined
Database Software	SQL Server 6.5
System	Unisys QS/2
Processor	400 Mhz Xeon
Txn/Sec Requirement	100

4. Measurement Basis for User Defined Workloads

Several measurements were taken to obtain processor timings and IO counts for various SQL types which were selected based on their generic relevance to an OLTP application environment. Each of the generic SQLs are described in the following sections along with citing example business cases where applicable.

4.1 Selects

Three Select examples were used: single table, and two and three table nested joins. Format for each example is:

- A generic SQL. Where applicable for the nested joins, the lowest table number represents the outer table; and the highest, the inner table.
- A description of the keys/indexes
- Access method used by the RDBMS
- The selectivity based on the results of the measurements
- Applicable business case examples

The cases are as follows:

- Single table selects

```
SELECT t1.a, t1.b
FROM t1
WHERE t1.a in (v1, ... , vn)
ORDER BY t1.b
```

Index description:

SQL Server:

PRIMARY, CLUSTERED KEY:

t1: b, c

NON-CLUSTERED INDEX:

t1: a

Access method: via t1.a

Example Business Cases:

Find how many orders for a given day and the products sold.

Find number of customers in a given area.

- Two table join

```
SELECT t1.a, t2.a, t2.b, t2.c
FROM t1, t2
WHERE
        t1.a IN (v1, ... , vn)  AND
        t2.a = t1.a
GROUP BY t1.a, t2.b, t2.c
ORDER BY t1.a, t2.c, t2.b
```

Index description:

SQL Server:

PRIMARY, CLUSTERED KEY:

t1: a

t2: a, b

Access method:

t1 is outer table

t1 is accessed via cluster key t1.a

t2 is accessed via cluster key t2:a,b

Selectivity:

For each row selected from t1, an average of four rows are selected from t2. The selectivity specified in the sizer is the number of rows selected from t1.

Example Business Cases:

Find all the suppliers that supply a specific product to determine the best price. There would be on the average four suppliers per product.

Find all the airlines that fly to a certain city and determine the best price.

- Three table join

```
SELECT t1.b, t3.b, t1.d
FROM t1, t2, t3
WHERE
        t2.b  = 'xyz' AND
        t2.a = t1.c AND
         t3.a = t1.a AND
 t1.b in (v1, ... , vn)
ORDER BY t3.b, t1.b
```

Index description:

SQL Server:

PRIMARY, CLUSTERED KEY

t1: a
t2: a
t3: a, c

NON-CLUSTERED INDEX

t1: a, b

Access method:

Nested table order (outer to inner): t1, t2, t3
t1 is accessed via non-clustered index t1:a,b
t2 is accessed via primary, clustered key t2:a
t3 is accessed via primary, clustered key t3:a

Selectivity:

For each value of t1.b (see the where clause), approximately 5 rows in t1 match the criterion. For each row selected from t1, a row in t2 is selected 20% of the time; for each row selected from t2, an average of 3.9 rows are selected from t3. The selectivity specified for the sizer is the number of values t1.b specified in the where clause.

Example Business Cases:

Find the length of time required to complete customer orders in a given market segment.
Find the actual arrival times compared to the scheduled arrival times of flights to a given city.

Note: Using the first example above, the measurement would consist of determining the status of orders placed on certain dates from a selected segment of the customer population. For the database used in the measurements, the customer segment chosen places about 20% of all of the orders, and each order consists of approximately 4 items on the average. In the sizer, the selectivity requested is equivalent to the number of orders of a specific kind.

4.2 Insert

Each transaction consisted of 10 inserts and was followed by one commit. A measurement consisted of 100 transactions.

4.3 Update

Each transaction consisted of 10 updates and was followed by one commit. A measurement consisted of 100 transactions.

4.4 Delete

Each transaction consisted of 10 deletes and was followed by one commit. A measurement consisted of 100 transactions.

5. Methodology

The goal in the sizer's design and development has been to provide a user friendly software tool that will facilitate assisting the customer in determining an optimal system configuration that will meet the customer's application and workload needs. This sizer is focused on determining system requirements for customer applications that will run on the Unisys Enterprise Servers, using the Microsoft Windows NT operating system.

The NT sizer was first developed as a means to quickly estimate server scalability for the TPC-C benchmark. Added to that capability was the ability to quickly compare TPC-C and TPC-D measurement results across vendors. This version allows the user to define a database, OLTP application, and workload which are used to estimate configuration requirements for NT Servers and for applications that use the SQL Server or the Oracle DBMS. The Sizer's capabilities will increase with each new version.

A major focus of activity in a development of this type of tool is the defining and conducting of certain key measurements whose results lend themselves well to estimating, i.e., predicting, the resource usage, and consequently, the resource requirement, for specified applications. This section discusses not only the nature of these measurements but also the application of these measurements to the prediction methods within the sizer. Discussion is scoped to the sizer's current capabilities: TPC-C and user defined OLTP workloads.

5.1 TPC-C Workloads

For a TPC-C workload, the intent is to determine the number of processors and amount of memory and mass storage required to support the transaction rate where the database size grows with the transaction rate.

The estimates were taken from TPC-C benchmark measurement results obtained from the Mission Viejo performance lab in addition to estimates based on additional measurements. The additional configuration values were obtained via curve fitting and prior knowledge of the behavior of SMP systems.

5.2 OLTP User Defined Workloads

Each transaction in an OLTP application consists of some mix of SQL inserts, deletes, updates, and selects. However, it is assumed that any query, i.e., SQL select must be of short duration to satisfy the business requirement of quick response times. Thus, each select to a single table or to a group of tables (nested join) should be via indexed access. Accordingly, three generic, indexed select statements were chosen to represent scope of queries in an OLTP environment. This sample is shown together with example business cases in section 4.

The measurements for these examples were taken on an Enterprise Server, 4 processor, 200 MHz system with 1 GB memory in a client/server computing environment. Additional measurements were conducted on the Aquanta XR/6 and the QS/2 systems. The resource measurements were of the server only in processing the request. We note that the equipment provided at the network and the client can significantly affect performance; however, the scope of the measurements was on the server only. Impact of the client workstation on performance is highly dependent on the equipment used and was not considered for these measurements. Calculations for network traffic were based on assumptions about the data traffic generated from each SQL.

The database used for the measurements was tuned, via indexing, etc., to run the candidate OLTP type SQLs. Two database sizes, 100 MB and 1 GB, were used for the purpose of verifying measurement results and methodology. For each measurement, the primary results used in developing parameter relationships were: processor usage, SQL logical IOs, SQL scans, and selectivity (for the queries).

For the SQL selects, several selectivity cases were run in order to determine a relationship between resource usage and selectivity. Also, each SQL was subjected to the SQL server "Show Plan"; results of Show Plan were then compared for similar SQLs to ensure that similar SQLs were always run in the same manner.

For the insert, delete, and update measurements, a sequence of 10 SQLs was executed before a Commit. The cost of a Commit was included in the data used. For deletes, an average of 4 records were noted deleted per delete SQL. Thus, the user of the sizer must take this into consideration when defining the application and workload.

From these measurements we were able to determine parameter relationships which allow us to predict system resource usage for similar SQLs defined by the user of the sizer.

The phenomenon of increased transaction service time due to an SMP environment was also factored into the configuration estimates.